Digital image Steganalysis based on Visual Attention and Deep Reinforcement Learning

Donghui Hu, Shengnan Zhou, Qiang Shen, Shuli Zheng, Zhongqiu Zhao, Yuqi Fan

***Abstract*—Recently, adaptive steganography methods have been developed to embed secret information with the minimal distortion of images. As the opposite art, steganalysis methods, especially some convolutional neural networks (CNNs) based steganalysis methods, have been proposed to detect whether an image is embedded with secret information or not. State-of-the-art steganography methods hide secret information in different regions of an image with different probabilities. However, most of current steganalysis methods extract the steganalysis features from different regions without discrimination, which reduce the performance of the current deep-learning-based steganalysis methods when attacking adaptive steganography methods. In this paper, we propose a new self-seeking steganalysis method based on visual attention and deep reinforcement learning to detect JPEG-based adaptive steganography. Firstly, a region is selected from the image by a visual attention method, and a continuous decision is then made to generate a summary region by reinforcement learning. Thereby, the deep learning model is guided to focus on these regions which are favorable to steganalysis and ignore those regions which are unfavorable. Finally, the quality of training set and the detection ability of steganalysis are improved by replacing the mis-classified training images with their corresponding summary regions. Experiments show that our method obtains competitive detection accuracy, compared with other state-of-the-art advanced detection methods.**

***Index Terms*—Adaptive steganalysis, Deep reinforcement learning, Convolutional neural network, Visual attention**

I. INTRODUCTION

IMAGE steganalysis is the art to detect steganography. Steganography algorithms are designed to embed secret information into various digital images by modifying pixels or frequency coefficients. Steganalysis technique is a countermeasure technique to resist steganography. Its essence is to attack the suspicious carrier, and to detect whether the carrier carries secret messages by analyzing its statistical characteristics. Early steganography algorithms are usually non-adaptive, such as LSB (Least Significant Bit) [1] in spatial domain and J-Steg [2], OutGuess [3], F5 [4], nsF5 [5], MB [6] in JPEG domain. In recent years, the rapidly developed adaptive steganography has retained more complex image statistical properties, and greatly improved the security of the steganography, which puts forward huge challenges to the steganalysis.

In the spatial domain, there are many adaptive steganography algorithms such as HUGO (Highly Undetectable steGO) [7], WOW (Wavelet Obtained Weights) [8], S-UNIWARD (Spatial-Universal Wavelet Relative Distortion) [9], HILL (High -pass, Low-pass, and Low-pass) [10], and MiPOD (Minimizing the Power of Optimal Detector) [11]. In JPEG domain, state-of-the-art adaptive steganography algorithms include UED [12], UERD [13], and J-UNIWARD [9] (JPEG domain version of UNIWARD). Most of these algorithms are designed under the framework of minimizing a distortion function, in which each pixel of an image unsuitable for embedding information is firstly assigned a low embedding cost. Then, the steganographic images are obtained by some coding techniques, such as STCs (Syndrome-Trellis Codes) [14]. The coding techniques can also be applied to 3D video, such as HEVC [15–17]. Besides, the steganography algorithms based on automatic generation model develop rapidly in recent years and has high capacity, security and reliability. Hu *et al*. [18] use DCGANs to generate stego images according to secret information, which is called the novel image SWE method. Duan *et al*. [19] propose a novel coverless image steganographic scheme based on a generative model.

Modern steganalysis methods are usually designed by training a classifier fed with carefully designed steganalysis features. Most of current steganalysis methods (detailed in Section 2) rely on designing more effective features to improve the detection performance for adaptive JPEG steganography. One of main characteristics of those steganography methods is that each region in an image has a specific loading capacity after adaptive steganography. Generally, the relatively flat regions are unsuitable for modification and embedding, and the regions with complex textures can effectively disguise the fact of the existence of secret information due to their complex statistical properties. Unlike in the spatial domain, the JPEG domain based steganography is mostly achieved by modifying the quantized DCT coefficient, and then this effect will be diffused into the corresponding 8×8 block of the spatial domain. Fig. 1 shows two stegos and the corresponding modifications embedded by the J-UNIWARD at different embedding rates. It can be clearly found that, whether in the JPEG domain or in the spatial domain, the modified pixels are concentrated in a certain area (i.e., the area with complex texture) instead of being evenly distributed. So the effectiveness of extracting features from various regions is quite different. However, most of current deep learning models are single-scale, which means to treat all pixels with the same scale, and therefore, it is not easy to capture the most favorable features for classification.

So, in this paper, we propose a novel Self-Seeking method which can automatically search regions more favorable for steganalysis in images without prior knowledge or human

Donghui Hu, Shengnan Zhou, Qiang, Shen Shuli Zheng, Zhongqiu Zhao and Yuqi Fan are with the School of Computer and Information, Hefei University of Technology, Hefei 230009, China (e-mail: hudh@hfut.edu.cn; sarahchoy@163.com; watashi_shenq@163.com; zsl251@163.com; z.zhao@hfut.edu.cn; yuqi.fan@hfut.edu.cn).

Yuqi Fan is also with the College of Department of Computer Science, University of Texas at Dallas, 800 W. Campbell Road, Richardson, TX 75080, USA (email: yuqi.fan@utdallas.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2900076, IEEE Access

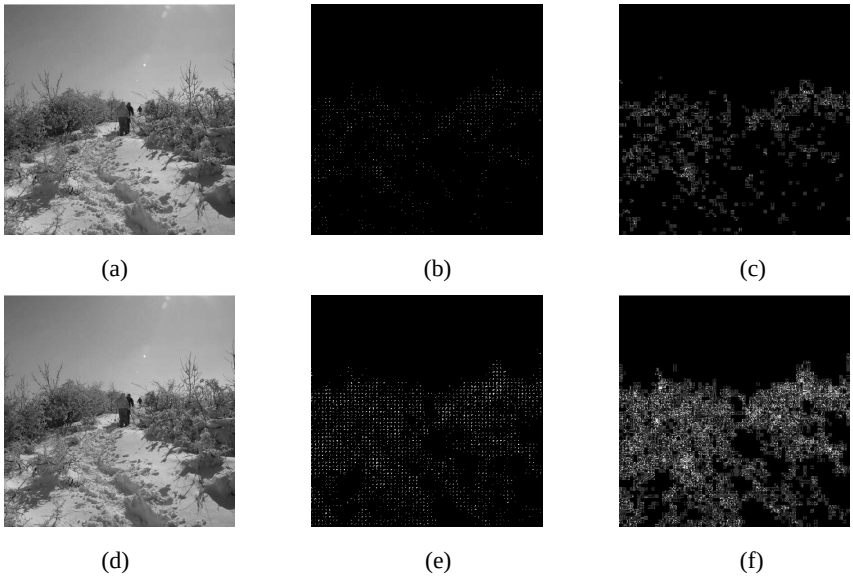


Fig. 1. The stegos and the corresponding modifications by the steganography algorithm J-UNIWARD at different embedding rates. (a) and (d) are the stegos with embedding rates of 0.1bpnzac and 0.4bpnzac, respectively. (b) and (e) are the corresponding modification points in DCT domain of images in Fig. 1(a) and (d), respectively. (c) and (f) are the modified points in the corresponding spatial domain of stegos in Fig. 1 (b) and (e), respectively. (Please note: to make the results more significant, we magnified the signal by 50 times.)

involvement. First, attention-focused regions (AFRs) are extracted from images by the visual attention model. Then the discriminant model (an existing stenanalysis method) judges whether the generated AFRs are benefit for steganalysis or not. The reward values are calculated based on the classification probabilities of AFRs by the top-level Softmax layer of the discriminant model. The summary of attention-focused regions (SoAFRs) is merged by continuous decision-making through reinforcement learning with the reward values. The redundant information of the features extracted from SoAFRs is less than that of the features extracted from original images. For a training set, we select the images which cannot be correctly classified by the the discriminant model, and replace them with the SoAFRs. Thereby, the designed SoAFRs can eliminate the regions which are unfavorable for steganalysis and thus improve the quality of the training set. Experimental results show that our method obtains competitive detection accuracy, compared with other state-of-the-art steganalysis methods.

The remainder of this paper is as follows. In Section 2, we review the state-of-the-art research related to our proposed method. In Section 3, we elaborate our method in detail. Experimental results and analysis are then given in Section 4. Finally, in Section 5, some conclusion are drawn.

II. Related work

A. JPEG steganalysis

With the development of image adaptive steganographic algorithms, the corresponding steganalysis techniques have been developed. The current mainstream steganalysis is mainly based on artificial feature extraction or deep learning based features. Fig. 2 presents the general flow diagram of two kinds of methods: the top one is based on hand-crafted feature sets and the bottom one is based on deep learning. Whether in

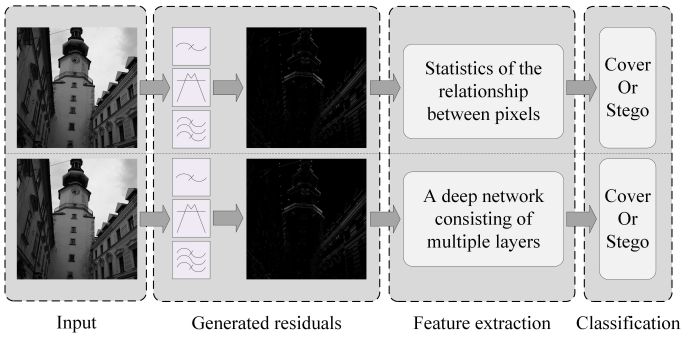


Fig. 2. Comparison of steganalysis methods based on hand-crafted feature sets and deep learning.

spatial or JPEG domain, the overall process of different steganalysis methods can be roughly divided into four steps while their main difference lies in the step of extracting features. The hand-crafted feature based methods usually extract features based on the relationship among pixels or DCT coefficients by statistical methods. While deep learning based methods generally design effective networks with different structures to automatically learn features.

Exploring content-adaptive schemes, some researchers have handcrafted various high-dimensional features, such as spatial rich model (SRM) [20–22], the selection-channel aware maxSRM and maxSRMd2 [23], and the steganalysis method [24] based on the IQM and the SRM in spatial domain. In JPEG domain, some modern schemes, such as DCTR (Discrete Cosine Transform Residual) [25], PHARM (Phase-Aware Projection Model) [26], GFR (Gabor Filter Residual) [27] and their variants [28, 29] extract features from the residuals of the decompressed JPEG images. These methods extract features in the spatial domain rather than in the JPEG domain, which seems more effective. Feng *et al*. [30] proposed

a novel scheme for constructing effective filters for JPEG steganalysis. As for binary image, Chen *et al*. [31] proposed a novel binary image steganalytic scheme, which is based on distortion level co-occurrence matrix.

Recently, with the development of deep learning, researchers have proposed various convolutional neural networks (CNNs) architectures [32] to extract features directly from input images. In spatial domain, Qian *et al*. [33] designed a CNN with image processing layer, convolutional layers and classification layers, obtaining the performance close to SRM. Further, Xu *et al*. in [34, 35] applied some strategies such as absolute value layer, batch normalization layer, and ReLU activation function, to the proposed network and achieved better performance than SRM. In [36], Ni *et al*. proposed a novel network structure which combined selection channel and adopted a new activation function called truncated linear unit (TLU). In [37], Hu *et al*. proposed a combined network consisting of three separate subnets with independent structures. The independent subnets can be repeatedly separated and merged and thus evolve with more diverse and effective features. Zhong *et al*. [38] introduced three ensemble methods aiming to increase the diversity between classifiers. In addition, some achievements of steganalysis based on deep learning has also been obtained in JPEG domain. In [39], Zeng *et al*. designed a large-scale hybrid deep-learning framework. The framework includes quantization and truncation phases which process residual noise to an interval, in order to accelerate the convergence on the basis of maximizing the retention of the original residuals. Considering the information loss in the processing of pooling, Xu *et al*. [40] replaced pooling layers with convolutional layers, and proposed a 20-layer deep residuals network to test J-UNIWARD.

Overall, most of CNN models treat all pixels with the same scale, which leads to the possibility of extracting unfavorable features. So in this paper, we use a visual attention method to address this problem.

B. Visual attention

When humans observe an image or understand a message, the brain receives the entire information. But at a given moment, the attention generally focuses on a part of the input information. And other parts of the information are eye-catching but allocated few resources. This kind of resource allocation is actually caused by the attention mechanism of the human brain. In recent years, the combination of visual attention mechanism and neural network has become a hot topic in deep learning. In 2014, Mnih *et al*. proposed the RAM model [41] (Recurrent Attention Model) to classify images, adding the attention mechanism to the traditional RNN, and using reinforcement learning to select the image position to be processed. In [42], Xiao *et al*. achieved the effectiveness of distinguishing subtle differences by filtering images twice. Liu *et al*. [43] proposed a fully Convolutional Attention Localization Network based on reinforcement learning to adaptively select multitasking driven visual attention regions, which can locate multiple blocks and enlarge the blocks to achieve fine grained recognition. Zhao *et al*. [44] applied attention to vehicle identification, and proposed a residual attention network that combines extremely deep convolutional neural networks with human visual attention mechanism. In [45], Wang *et al*. proposed Residual Attention Network, a convolutional neural network incorporating attention mechanism in an end-to-end training fashion. From Fig. 1, we find that the JPEG adaptive steganography algorithm can adaptively select positions with less distortion, and the modified points are mainly concentrated in some certain regions. Thereby, in this paper, we introduce the attention mechanism into the deep learning based steganalysis, aiming to enable the network to concentrate on those regions favorable for the steganalysis. It is a sequential decision process to choose the AFR by using attention mechanism, and we use reinforcement learning to deal with this problem.

C. Reinforcement learning

Reinforcement learning [46] is an important branch of machine learning, which can effectively solve sequential decision problems. It learns how to achieve goals in a complex and uncertain environment. Reinforcement learning is widely applied in many areas, including controlling robots, managing merchandise inventory, and playing game. It can adapt to the changing environment and response with a series of corresponding actions to approach ultimate goals. For example, reinforcement machine learning is adopted to enhance the accuracy of an ensemble system consisting of multiple feature extractors and multiple classifiers (MFMC) [47], which is applied to detect pedestrian and to recognize handwritten numerals. AlphaGo [48, 49], a game program that has recently shown extraordinary talents in the international arena, needs to learn the most favorable tactics for victory in various situations. In this system, the basic components of reinforcement learning include external environment, learner agent, action space, reward function, policy, etc. Agent interacts with the external environment in the process of learning. The interaction process is completed as follows: the agent selects an action at a certain moment, the environment gives a reward (or punishment) according to the reward function, and then inform it of the next state. After this cycle, the agent learns a series of optimization strategies. The essence of reinforcement learning is a Markov decision process. The ultimate goal is to maximize the overall reward in the decision making process to achieve the desired optimum.

III. The Self-Seeking Method

The state-of-the-art adaptive steganography methods usually first assign a distortion value to each pixel via a distortion function based on the embedding cost, and then some advanced coding techniques, such as STCs, will be applied to minimize the expected distortion value for all pixels in texture areas. Obviously, the amount of information that can be carried by different regions is quite different. In many traditional methods [23, 28, 29], it has been proved that the validity of the extracted features in different regions is inconsistent. At present, most CNNs used for steganalysis treat all pixels at a single scale, which is unable to effectively capture the key

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2900076, IEEE Access

points or areas more favorable to steganographic detection. Besides, it may also cause the information redundancy, and further affects the classification accuracy. Taking it into account, we design the Self-Seeking method to find the favorable areas automatically.

A. The Architecture of Self-Seeking method

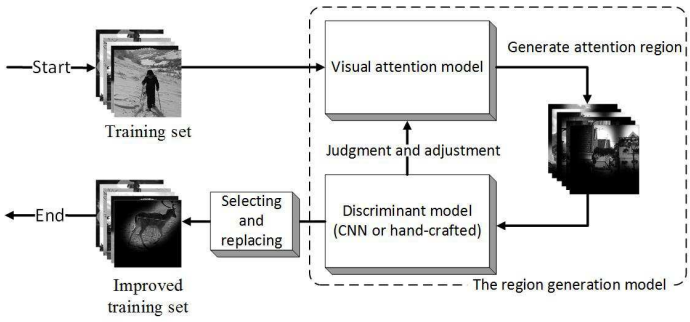


Fig. 3. The framework of self-seeking method.

Our approach is based on visual attention mechanism and reinforcement learning. The attention mechanism is to focus on a selected region with "high resolution", and to use "low resolution" to perceive the surrounding pixels roughly. After the AFR is generated by this mechanism, the region enters our "brain" for comprehension evaluation. With the feedback of the "brain", the location of the focused point is constantly adjusted by our "minds", and thereby a SoAFR is finally generated.

Fig. 3 shows the overall structure of our proposed method based on visual attention and reinforcement learning. It mainly consists of four parts: the discriminant model for "brain", the visual attention model for "eyes", the region generation model for recursive evaluation for "minds", and the selecting and replacing model. We input the images from training set into visual attention models one by one to generate AFRs. Subsequently, the focus of eyes are constantly adjusted through the judgment of the discriminant model, and thereby the SoAFRs are generated.

The core part is the region generation model, which is running on the basis of the visual attention model and discriminant model, and the reinforcement learning is applied into this model to form the SoAFRs. Fig. 4 shows the process of generating the SoAFRs. In our designed reinforcement learning, there are 16 actions, which run repeatedly 4 times. Each AFR is generated by the visual attention model, and then is inputted to the discriminant model to obtain a score as the reward of the reinforcement learning. Continuous decision-making through reinforcement learning are made in each iteration to output an AFR with the maximal reward. The final SoAFRs are formed by merging the 4 selected AFRs.

Each part of the mentioned process will be described in detail below. After this process, we hope to use the SoAFRs to update the training set so that the discriminant model may not learn negative features. To avoid losing some useful information in abandoned regions, we adopt the selecting and replacing model to refining the training set by filter the regions that meet the requirements which are set in Section 3.5.

B. The visual attention model

In the field of computer vision, attention mechanism can be realized in various forms, which can be roughly divided into soft attention and hard attention. The typical examples of soft attention are Residual Attention Network [50] and Two-level Attention [42]. Soft attention takes care of all positions at a time, but the weights of different positions are different. The attention of this mechanism is relatively divergent and can be trained through back propagation. The main equation of soft attention is

$$Z_t = \sum_{i=1}^{L} \alpha_{t_i} \times a_i, \tag{1}$$

where a_i is the ith D-dimensional vector of an image, and α_{t_i} is the weight of a_i. Hard attention focus on specific areas, which usually uses reinforcement learning to predict the areas. And the attention mechanism used in this paper belongs to hard attention.

When the neural network extracts the features of an image, we can design an attention model to reduce the interference of other factors. That is to say, the region from which the features are extracted is displayed explicitly, while other regions which are farther away from the focused area are blurred or not displayed. Suppose the size of the image is $n_1 \times n_2$, and the focused region can be calculated by

$$\mathbf{I}_t = \mathbf{I} \cdot \mathbf{K}, \mathbf{K} = (k(i,j)), \tag{2}$$

where $\mathbf{K}$ is the attention intensity matrix corresponding to each pixel, each element of which represents the degree of attention on a pixel, and $i \in \{1, ..., n1\}$, $j \in \{1, ..., n2\}$. The attention intensity means that each pixel in the image is assigned with a different level of attention, and the range of attention intensity is [0,1]. The model will give a pixel p a hundred percent attention intensity at first. The intensity values of the surrounding pixels will gradually decrease as their distances from the pixel p increase. The mathematical description of the attention intensity K is as follows:

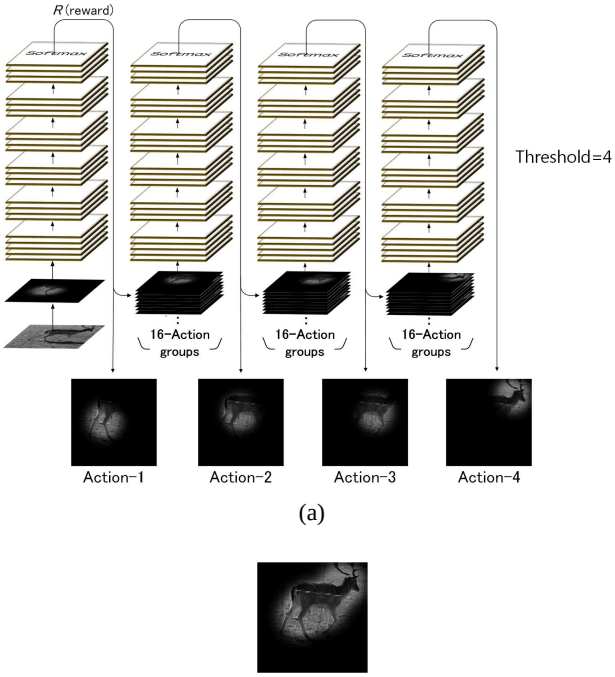
$$K(a,b) = 1 - \frac{1}{1 + exp(l_1 d + l_2)}, \tag{3}$$

where the parameters l_1 and l_2 are used to control the shape and size of the region, respectively, and d is the Euclidean distance between another pixel $p'(a,b)$ and the current attention pixel $p(i,j)$. The formula to calculate d is as follows:

$$d = \sqrt{(a-i)^2 + (b-j)^2}. \tag{4}$$

Eqs. (3) and (4) together constitute the mathematical expression of the visual attention mechanism. For more intuitive observation, we visualize this expression in Fig. 5. In this case, we set l_1 and l_2 to be -0.06 and 6, respectively. The scope of the AFR is a circle with a radius of 200 pixels. We defined the radius of the circle as r, so in this case, radius $r = 200$. We study the influence of different parameters in the later experimental part, so as to select a more suitable region radius. After calculating the attention intensity of all pixels,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2900076, IEEE Access



(a)

(b)

Fig. 4. The process of generating the attention region based on visual attention mechanism and reinforcement learning. (a) The overall structure of the model, the threshold is set to 4 and there are 16 candidate actions. (b) The summarized result of multiple regions selected by the model.

the corresponding AFR can be generated. Fig. 6 is an example of the results calculated from this visual attention model.

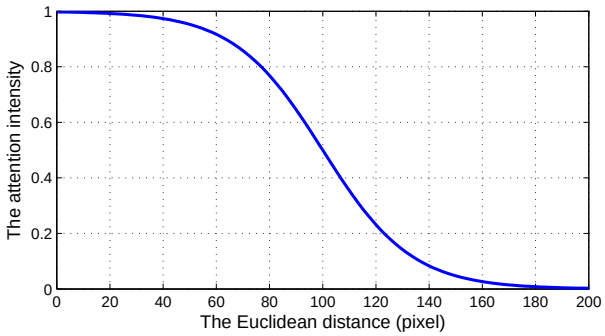
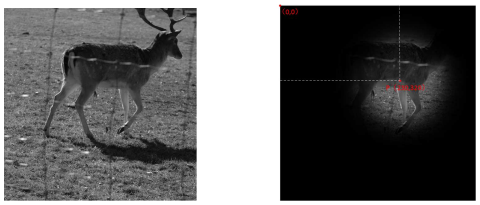


Fig. 5. The mathematical expression of visual attention mechanism. The attention intensity decrease when the distance from the attention pixel increases.



(a) (b)

Fig. 6. An example of visual attention model. (a) The original image. (b) The generated AFR when the attention point is located at the point $p(230, 320)$ in the image.

In fact, it is not clear whether the obtained AFR is effective for the classification or not. Therefore, it is necessary to divert attention points in order to ensure the region is conducive to steganalysis. We design a discriminant model to evaluate the region to determine whether it meets the requirements or not. The transfer of attention will continue according to the regional evaluation. The transferring route is actually a process of sequential decision, which is a decision-making method for optimization of stochastic or uncertain dynamic systems. We use the reinforcement learning to conduct subsequent selection of attention regions, the reward of which is calculated by the discriminant model.

C. The discriminant model

The discriminant model is used to evaluate whether a region meets the requirements and to guide the shift of sight. It can be a CNN or hand-crafted based steganalysis classifier. In this paper, we use the former one.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2900076, IEEE Access

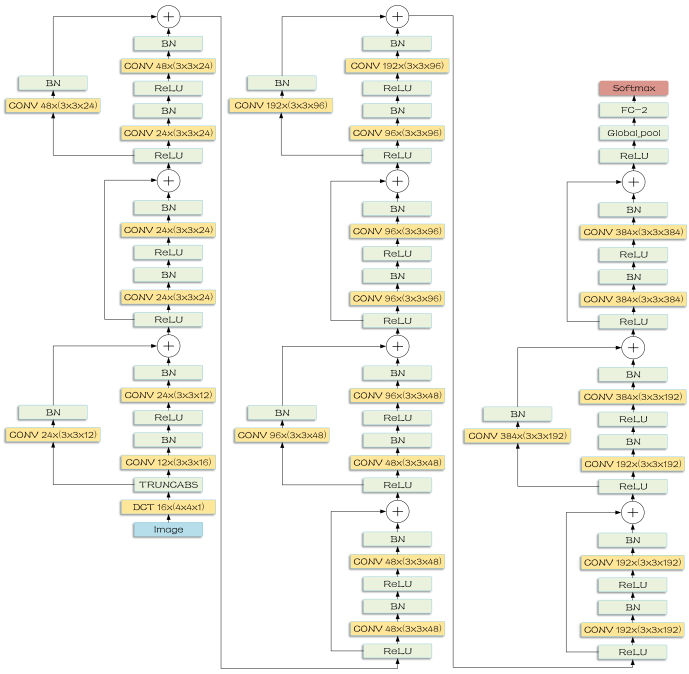


Fig. 7. The structure of the discriminant model, a deep residual network.

Fig. 7 shows the details of the discriminant model provided by [40], which is a deep residual network. The JPEG image is firstly transformed into the spatial domain, and then enter the pre-processing layer containing 16 DCT kernels for noise residual extraction. The network consists of 25 convolutional layers, multiple activation layers, batch normalization layers, a global pooling layer, and a truncated layer (TRUNCABS).

D. The region generation model based on reinforcement learning

The region generation model involves 3 inputs and 1 output. The 3 inputs are as follows: (1) S, which means the state of a specific moment; (2) A, which means a set of subsequent actions that can be taken in a state; and (3) R, which represents the reward or punishment value that comes into each state. The output of the model is the policy π, which is the action sequence selected according to different rewards in the state S. To gradually adjust the observation position to approach the target state under the influence of the environment, we need to choose an initial location that is suitable for steganalysis and further to optimize with this initial location. The mathematical solution can be described as:

$$\pi(S') = argmax_a R(S', \mathbf{A}), \tag{5}$$

where $a \in A$, R and S' are the action, reward function and current state, respectively. Eq. (5) means that selecting a in the action set $\mathbf{A}$ under current state S' will get the maximum reward value, which allows one to get the best possible transfering route of actions to approach the ultimate goal–a region that is beneficial for steganalysis.

1) States: The state is the focus point (or the center) of the current AFR (for example, the point $p(230, 320)$ in Fig. 6(b)). In the first run, the model randomly selects a focus point to initialize the state. And then, the model iteratively updates the current state according to the feedback from the external environment.

2) Actions: When eyes are observing a picture and concentrating on a certain region, they generally expand along the region in some directions. Different directions or distances may get inconsistent and unpredictable information. So we design a set of follow-up 16 actions consisting of 8 directions and 2 distances. Assuming that the attention point is $p(i_s, j_s)$ at the state s, then the attention point $p(i_{s+1}, j_{s+1})$ at the next state is defined as follows:

$$p(i_{s+1}, j_{s+1}) = f(p(i_s, j_s), \theta, \psi), \tag{6}$$

where θ, ψ are the direction angle and the distance magnitude of possible movement, respectively. Fig. 8 shows 16 transferring actions starting from the attention point in Fig. 6(b). We choose 8 directions as being the forward routes of travel: upper, lower, left, right, upper left ($45°$), top right, lower left and bottom right. In each direction, there are two distances of d_1 and d_2 pixels respectively. The 16 points by 16 actions generate the next 16 attention regions. Then, we need to select one from these 16 regions which is most favorable to our classification, and then recursively update the attention point starting from the new attention region.

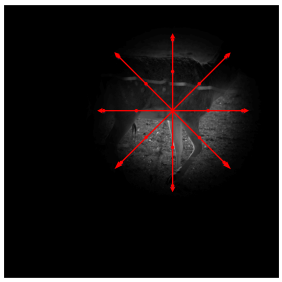


Fig. 8. The simulation diagram of 16 kinds of visual transferring routes.

3) Rewards: The attention regions are inputted to the discriminant network. Thereby, we can obtain the hierarchical expression of information at each layer of the network, get the classification probabilities of the regions at the top-level Softmax layer, and achieve the corresponding reward values. The Softmax is calculated as follows:

$$\phi_i(z) = \frac{e^{z_i}}{\sum\limits_{j \in group} e^{z_j}}, \tag{7}$$

where the number of group is 2 (one group is cover, and the other group is stego), z_i and z_j are the ith and jth output of the discriminant model, respectively. The Softmax is the probability distribution of each class and the sum of all neuron outputs is 1.0. When the output of a signal source is uncertain, the determination can be measured by the probability of occurrence of various categories. The greater the probability is, the bigger the number of occurrences is and the smaller the uncertainty is. For example, if the discriminant model determines that the probability of belonging to class A (such as stego) is 0.75 and the probability of belonging to class B (such as cover) is 0.25, the model will assign the image to class A. In the case of the groundtruth class label being A, though

Algorithm 1 RL-based algorithm for generating a SoAFR.

Input: An input image I.
Output: The region O, save as JPEG format.

Initialize the current status s with a random point p;
Initialize the region O with random weights;
for episode=1,T **do**
    Calculate 16 subsequent actions **A** according to s using Eq. (8);
    Generate 16 focused images $\mathbf{I}_t$ according to **A** using Eqs. (2), (3), and (4);
    Calculate 16 rewards **R** according to $\mathbf{I}_t$ using Eqs. (7), (8), and (9);
    if Fall into local optimal **then**
        Initialize the current status s with a random point p';
        continue;
    else
        Execute the action a_{max} in **A** which has the max R in **R**;
        Update the current status s according to a_{max};
        Update all parameters in Eq. (5);
        Store the current status s into a matrix **P**;
Generate SoAFR O from all s in **P**;
Return SoAFR O as a JPEG format image;

the classification result is correct, it is not stable enough. If the model determines that the probability of belonging to class A is 0.95, the probability of belonging to B is 0.05, the image may have stronger discriminative features. Based on this characteristic, we calculate the information entropy of the probability distribution of output to measure the certainty of whether the region is cover or stego. Assuming that the signal source outputs classes U_1 and U_2 with the corresponding probability P_1 and P_2, the mathematical description of the information entropy is as follows:

$$H(U) = E[-logP_i] = -\sum_{i=1}^{2} P_i logP_i. \tag{8}$$

The greater the uncertainty is, the larger the entropy is. Therefore, the attention region with lower entropy is more effective for steganalysis. The "rewards and punishment" strategy is designed as follows:

$$R = \begin{cases} 0 & \hat{\xi} \neq \xi \\ 1 & \hat{\xi} = \xi, H(U_i) \neq minH(U), \\ 2 & \hat{\xi} = \xi, H(U_i) = minH(U) \end{cases} \tag{9}$$

where $\hat{\xi}$ is the predicted label of the region and ξ is the actual label.

4) Prevent from falling into local optimum: If the initial attention point is not well selected, the searching is easy to fall into a local optimum. In this case, we design a solution to artificially transfer routes when the 16 selected regions are not satisfactory: the worst case is that all their predictions are wrong. With the memory of the focused points selected before, at this time the model reselects another focused point. Algorithm 1 summarizes the calculation and extraction process of SoAFR. When an image is inputted into the model, a series of AFRs will be obtained. The model records all the regions which are favorable to steganalysis and merge them into an SoAFR. All SoAFRs will be used to train a new discriminant model.

E. Selecting and replacing model

Since our method needs to know the real label information of an image first, we only use it to the training set to improve the effectiveness of the training set. The main idea is to replace part of the images which cannot be recognized correctly with their corresponding SoAFRs.

To verify the validity of the extracted regions, we propose to replace all images in the training set with SoAFRs. The average file size of the cover images in BOSSbase 1.01 [51] (the quality factor is 75, and the stegos are embedded by J-UNIWARD with embedding rate of 0.4bpnzAC) is 32222B≈31.47KB, and the average file size of the corresponding stegos is 32264B≈31.51KB. However, the average file size of the cover set consisting of SoAFR is only 19.34KB, and the average file size of the stego is 15.45KB. Therefore, replacing all images can cause massive information loss. So we replace only a portion of the training images with SoAFR. In general, the images which cannot be recognized correctly may carry more interference features unfavorable to steganalysis, while the SoAFRs of these images may exclude some unfavorable features. Therefore, we replace the images which cannot be correctly distinguished by the discriminant model with their corresponding SoAFRs. The file size of the SoAFR is also important. If the SoAFR is very small, it may lose a lot of information valuable for steganalysis; if the file size of SoAFR is very big and close to that of the original image, it may has some redundant information that unfavourable to steganalysis. We will validate this in the experiments and analysis part.

Algorithm 2 shows the main steps of the proposed selecting and replacing model to improve the training set. One image which cannot be distinguished by the discriminant model correctly, the image can be replaced by its SoAFR if the following conditions are satisfied: the file size of the its SoAFR is greater than T_1, and the distance between the file sizes of the image and its SoAFR is smaller than T_2.

IV. Experimental results and analysis

A. Dataset and Settings

The image dataset used in the experiments originates from BOSSbase v1.01, which contains 10,000 grayscale images of size 512×512. The adaptive steganography method in this paper is J-UNIWARD, and the default parameters are used during embedding. For comparison, we use 17000-dimensional SCA-GFR [28] as the traditional artificial feature-based method, and use Xu-CNN [40] as the deep learning based method. The number of each data set is consistent with that in [40]. We use two high performance graphics cards, NVIDIA Geforce GTX TITAN X and NVIDIA Quadro K5200, to speed up the computation and optimization. We set the initial learning rate of the deep learning model to 0.001, which is a one-tenth reduction in every 5,000 iterations, and the maximum

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2900076, IEEE Access

Algorithm 2 The selection and replacing alogrithm.

Input: The training set **S**, the error image set **I** which is none, the thresholds T_1 and T_2 that regular the file size of the SoAFR.

Output: The replaced image set (training set) **O**, save as JPEG format.

for S : **S do**
  //traverse all elements S in the set **S**
  if the image S cannot be distinguished by the discriminant model correctly **then**
    Put S into the error image set **I**, $\mathbf{I} = \mathbf{I} \bigcup S$.
for I : **I do**
  //traverse all elements I in the set **I**
  Calculate the SoAFR of I and denote it as I_a
  Get the file sizes of I_a and I and denote them as S_I^a and S_I, respectively.
  if $S_I^a > T_1$ and S_I-$S_I^a < T_2$ **then**
    Update $I \leftarrow I_a$;
Return **I**;

number of iterations was set to 120,000. The momentum value and weight decay for gradient descent are set to 0.9 and 0.0005, respectively. In the preprocessing layer of the model, the value of the high pass filter kernel is fixed when the noise residuals are extracted, and the back propagation and updating parameters are not needed. The learning rate and the weight decay of this layer are all set to 0. We first update and improve the quality of the training set through our method, and use the model [40] as the discriminant model. We choose the classification error rate as the evaluation index.

B. The effectiveness of the method

In this subsection, we show the selected SoAFRs extracted by our method. In addition, we verify the necessity of only replacing part of training image sets with SoAFRs. There are mainly several variables that can be controlled by humans: the size of the attention region (radius, r), and the number of iterations of reinforcement learning (threshold, T). We then compare the effects of different parameters through experiments and give the final results using a set of competitive parameters.

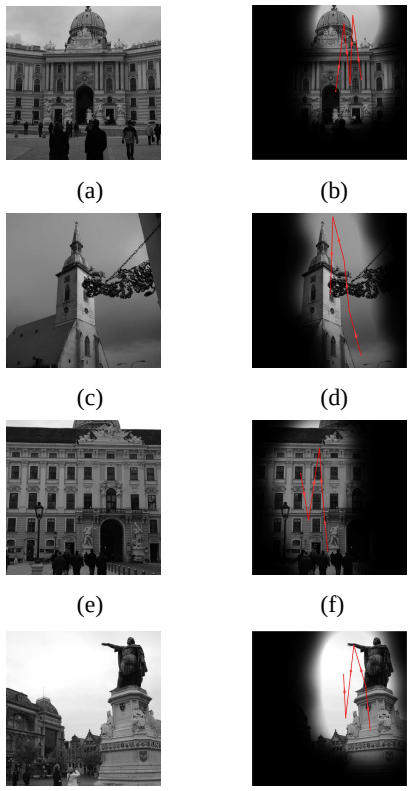


Fig. 9. The transfer routes of the visual attention points. (a),(c),(e) and (g) are the original images from BOSSbase v1.01; (b),(d),(f) and (h) are the attention regions corresponding to images (a),(c),(e) and (g), respectively.

1) Generating region: Fig. 9 shows the transfer routes of visual attention and the SoAFR generated by the the proposed Algorithm 1, with radius $r = 200$, iterations $T = 5$, the embedding rate 0.4bpnzac and the quality factor 75. Note that the subsequent experiments are subject to this condition unless otherwise specified. Among them, (a), (c), (e) and (g) are randomly selected JPEG images in the training set, while (b),(d),(f) and (h) are corresponding SoAFRs. The lines in images (b),(d),(f) and (h) represent the specific routes of visual attention transfer, and the arrow represents the direction of transfer. Fig. 10 shows the "jumping" (discontinuous) SoAFR generated by introducing corrective measures described in subsection III-D4 which prevents the model from falling into a local optimum. The following rules can be found by generating different types of SoAFR images through experiments.

a) For a given image, only a part of the image generated by the visual attention method is bright (or selected), and the size of the SoAFR or the amount of information it carries is smaller than that of the original image.
b) Visual attention points move along the texture edges of the images.
c) Regardless of where the initial point of attention is, the final transfer direction is the region that is moderate toward the amount of embedded information, that is, an region that is not particularly smooth or particularly texture complex.
d) The SoAFR may be discontinuous.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2900076, IEEE Access

(a) (b)

Fig. 10. The examples of discontinuous visual attention areas.

There are probably several reasons for above-mentioned phenomena. First, due to the characteristics of adaptive steganography, the particularly smooth region is not favorable to the extraction of effective features because of the small amount of embedded information, and the complex texture regions with a large amount of embedded information are also relatively difficult to be detected because of the complex statistical characteristics. The SoAFRs extracted in this paper means that perhaps the region with a moderate amount of embedded information (or moderate complex of texture) is more suitable for steganalysis. Secondly, the region in the image that facilitates steganalysis may not be continuous. When an image has more than one complex texture region, then if only one continuous region is extracted, it is easy to fall into the local optimum and thus reduce the effectiveness of final steganalysis. In this case, skipping the inappropriate regions (i.e., smooth regions) can effectively alleviate the phenomenon.

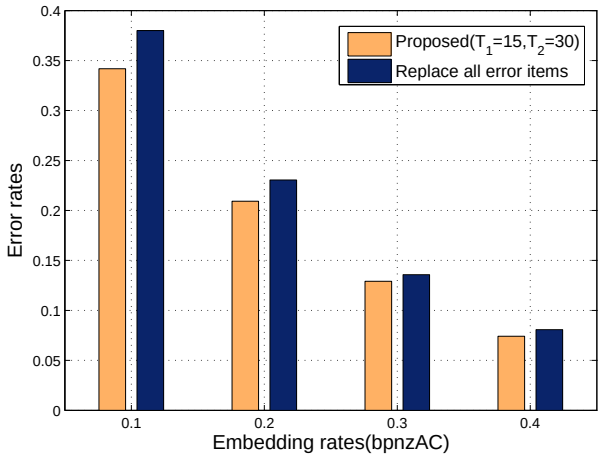


Fig. 11. The comparison results of performance when using different replacing strategies.

2) Selecting appropriate region: In order to further verify the validity of the extracted region images, we replace the training images with their corresponding SoAFRs. We first replace all images in the training set with their corresponding SoAFRs (the first replacement strategy). After training, we found that both the model learning efficiency and the detection error rate are very poor. The detection error is 0.4422 when the embedding rate is 0.4bpnzac and the quality factor is 75. The main reason (we have described in subsection III-E) is that the attention image loses nearly half of the information of the image, making the image lose many important features during the training and learning process. So we change the replacement strategy to only replace the images that are misclassified in the training set with their corresponding SoAFRs (the second replacement strategy). The other replacement strategy is to replace the misclassified images with only the region with the file size larger than 15KB and less than 30KB compared with that of the original image (the third replacement strategy), which is proposed in Algorithm 2, with the parameters $T_1 = 15$ and $T_2 = 30$. The reason for using these values is that the average values of the SoAFRs and original images are about 15KB and 30KB, respectively. In the real-world applications, we can set the values of T_1 and T_2 based on the actual file size of the detected images and generated SoAFRs, and adjust them constantly according to the experimental results.

Fig. 11 shows the comparison results of the second and the third replacement strategies when the JPEG quality factor is 75, and the steganography algorithm is J-UNIWARD with the embedding rates range from 0.1 bpnzAC to 0.4 bpnzAC. It can be found that the detection performance of the third replacement strategy is apparently better than that of the second replacement strategy, which proves the effectiveness of the proposed the Algorithm 2.

3) Setting parameters: There are a few other human-controllable variables in reinforcement learning, such as the size of the attention region (radius, r), and the number of iterations of reinforcements (T in Algorithm 1). We do two sets of experiments to compare the detail performances of the proposed method with different parameters.

In the first set of experiments, we set the radius r with constant value of 200, while the number of iterations T are 3, 5 and 7. The experimental results are shown in Table I. In general, increasing the iterations of reinforcement learning means that the model consumes more time and may get more information carried in the extracted regions. However, if the iteration of reinforcement learning is too large, the extracted regions will contains more unfavorable features to steganalysis which leads to the decline in detection accuracy. From the experimental results we can see the number of iterations has a subtle effect on the detection performance. When the number of iterations is 5, the detection result of the model has a slight detection advantage (with the lowest error rate of 0.0807).

TABLE I
THE COMPARISON RESULTS OF PERFORMANCE UNDER DIFFERENT ITERATIONS.

Algorithm	$T = 3$	$T = 5$	$T = 7$
J-UNIWARD	0.0818	0.0807	0.0878

In the second set of experiments, we set the number of iterations T with constant value of 5, while the radii rs are 100, 200 and 300. The experimental results are shown in Table II. Comparatively, when r=200, the detection result is the best in the three cases. The reason is that when the visual attention area is very small, a large amount of information is lost in the model, which is obviously unfavorable for steganalysis. As the size of the visual attention region increases, more valuable information is kept. However, when the size of the

visual attention region exceeds a certain value, the amount of redundant and even unfavourable information is also kept and finally decreases the performance of the proposed method.

TABLE II
THE COMPARISON RESULTS OF PERFORMANCE UNDER DIFFERENT RADII.

Algorithm	$r = 100$	$r = 200$	$r = 300$
J-UNIWARD	0.0847	0.0807	0.0842

TABLE III
CLASSIFICATION ERRORS FOR DIFFERENT EMBEDDING RATES.

Algorithm	QF	bpnzAC	Proposed	Xu	SCA-GFR
J-UNIWARD	75	0.1	**0.3418**	0.3454	0.3588
		0.2	**0.2093**	0.2147	0.2301
		0.3	**0.1292**	0.1335	0.1401
		0.4	**0.0741**	0.0794	0.0803
	95	0.1	0.5000	0.5000	**0.4640**
		0.2	0.4066	0.4134	**0.4017**
		0.3	**0.3229**	0.3258	0.3347
		0.4	**0.2531**	0.2586	0.2633

4) Final results: Through the above experiments, the parameters that are competitive in all cases are selected: $r = 200$, $T = 5$, $T_1 = 15$, $T_2 = 30$ when $QF = 75$, $T_2 = 60$ when $QF = 95$. We do comparative experiments of our method using above parameters with Xu's method [40] and SCA-GFR[28]. The comparison results are shown in Table III, from where we can see that our method outperforms Xu's method and SCA-GFR in detection accuracy. For the J-UNIWARD with low embedding rates of 0.2∼0.4 bpnzac when QF is 75 and 0.2bpnzac and 0.4 bpnzac when QF is 95, our method effectively outperforms Xu's method by 0.5%∼1%. Note that in the experiments, our method uses Xu's method [40] as the discriminant model. Although compared with Xu's method [40], the detection accuracy of our method is improved, the detection results of our method still depend on the performance of the discriminant model.

Fig. 12 shows the comparison results between our method and Xu's method [40] when the embedding rate is 0.4bpnzAC. We can find that as the number of iterations increases, the error rates of our method decrease fast and tend to be stable in little time. We also can see the error rates of our method are lower than that of Xu's method in each training iterations.

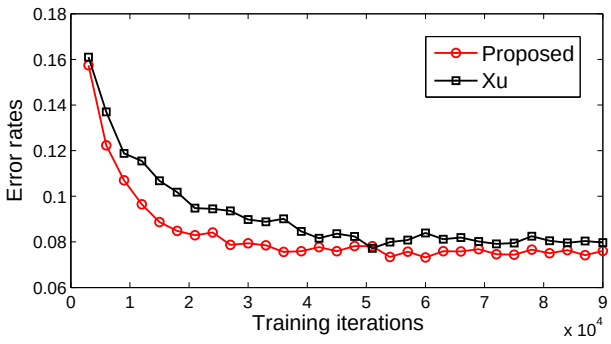


Fig. 12. Comparison of effectiveness of the model at 0.4bpnzAC.

V. CONCLUSIONS AND FUTURE WORK

This paper proposes a deep learning steganalysis method based on visual attention and reinforcement learning, aiming at the characteristics of the adaptive steganography algorithm for JPEG images. Our method converts the image into AFRs by a visual attention model, and then makes continuous decision by reinforcement learning to select SoAFRs which are more favorable to steganalysis. Experimental results show that our proposed method can effectively improve the quality of training sets, and eliminate the unfavorable features in training process, and finally improve the detection accuracy of steganalysis.

Due to time restriction, we only chose Xu's model as the discriminant model in this paper. However, the discriminant model can be replaced by other base steganalysis methods. in the future, we will research on new steganalysis method based on other types of attention mechanism and reinforcement learning, and try other discriminant models.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (NSFC) under the grant No. U1836102, and the Natural Science Research Project of Colleges and Universities in Anhui Province under the grant No. KJ2017A734.

REFERENCES

[1] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding-a survey," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1062–1078, 1999.

[2] D. Upham, "Steganographic algorithm JSteg," *Software available at http://zooid. org/˜ paul/crypto/jsteg*, 1993.

[3] N. Provos, "Defending against statistical steganalysis." in *Usenix security symposium*, vol. 10, 2001, pp. 323–336.

[4] A. Westfeld, "F5-a steganographic algorithm," in *International workshop on information hiding*. Springer, 2001, pp. 289–302.

[5] J. Fridrich, T. Pevnỳ, and J. Kodovskỳ, "Statistically undetectable JPEG steganography: dead ends challenges, and opportunities," in *Proceedings of the 9th workshop on Multimedia & security*. ACM, 2007, pp. 3–14.

[6] P. Sallee, "Model-based steganography," in *International Workshop on Digital Watermarking*. Springer, 2003, pp. 154–167.

[7] T. Pevnỳ, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *International Workshop on Information Hiding*. Springer, 2010, pp. 161–177.

[8] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on*. IEEE, 2012, pp. 234–239.

[9] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, no. 1, p. 1, 2014.

[10] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Image*

Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014, pp. 4206–4210.

[11] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 221–234, 2016.

[12] L. Guo, J. Ni, and Y. Q. Shi, "An efficient JPEG steganographic scheme using uniform embedding," in *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on*. IEEE, 2012, pp. 169–174.

[13] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, "Using statistical image model for jpeg steganography: uniform embedding revisited," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2669–2680, 2015.

[14] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 920–935, 2011.

[15] Z. Pan, X. Yi, and L. Chen, "Motion and disparity vectors early determination for texture video in 3d-hevc," *Multimedia Tools and Applications*, pp. 1–18, 2018.

[16] Z. Pan, J. Lei, Y. Zhang, and F. L. Wang, "Adaptive fractional-pixel motion estimation skipped algorithm for efficient HEVC motion estimation," *TOMCCAP*, vol. 14, no. 1, pp. 12:1–12:19, 2018. [Online]. Available: https://doi.org/10.1145/3159170

[17] J. Lei, J. Duan, W. Feng, N. Ling, and C. Hou, "Fast mode decision based on grayscale similarity and inter-view correlation for depth map coding in 3D-HEVC," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 28, no. 3, pp. 706–718, 2018.

[18] D. Hu, L. Wang, W. Jiang, S. Zheng, and B. Li, "A novel image steganography method via deep convolutional generative adversarial networks," *IEEE Access*, vol. 6, pp. 38 303–38 314, 2018.

[19] X. Duan, H. Song, C. Qin, and M. K. Khan, "Coverless steganography for digital images based on a generative model," *Computers, Materials & Continua*, vol. 55, no. 3, pp. 483–493, 2018.

[20] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics & Security*, vol. 7, no. 3, pp. 868–882, 2012.

[21] W. Tang, H. Li, W. Luo, and J. Huang, "Adaptive steganalysis against WOW embedding algorithm," in *ACM Workshop on Information Hiding and Multimedia Security*, 2014, pp. 91–96.

[22] ——, "Adaptive steganalysis based on embedding probabilities of pixels," *IEEE Transactions on Information Forensics & Security*, vol. 11, no. 4, pp. 734–745, 2016.

[23] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for steganalysis of digital images," in *IEEE International Workshop on Information Forensics and Security*, 2015, pp. 48–53.

[24] Y. Yang, Y. Chen, Y. Chen, and W. Bi, "A novel universal steganalysis algorithm based on the IQM and the SRM," *Computers, Materials & Continua*, vol. 56, no. 2, pp. 261–272, 2018.

[25] V. Holub and J. Fridrich, "Low-complexity features for JPEG steganalysis using undecimated DCT," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 2, pp. 219–228, 2015.

[26] ——, "Phase-aware projection model for steganalysis of JPEG images," in *Media Watermarking, Security, and Forensics 2015*, vol. 9409. International Society for Optics and Photonics, 2015, p. 94090T.

[27] X. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang, "Steganalysis of adaptive JPEG steganography using 2D Gabor filters," in *Proceedings of the 3rd ACM workshop on information hiding and multimedia security*. ACM, 2015, pp. 15–23.

[28] T. D. Denemark, M. Boroumand, and J. Fridrich, "Steganalysis features for content-adaptive JPEG steganography," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1736–1746, 2016.

[29] C. Xia, Q. Guan, X. Zhao, Z. Xu, and Y. Ma, "Improving GFR steganalysis features by using Gabor symmetry and weighted histograms," in *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2017, pp. 55–66.

[30] G. Feng, X. Zhang, Y. Ren, Z. Qian, and S. Li, "Diversity-based cascade filters for jpeg steganalysis," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2019.

[31] J. Chen, W. Lu, Y. Yeung, Y. Xue, X. Liu, C. Lin, and Y. Zhang, "Binary image steganalysis based on distortion level co-occurrence matrix," *Computers, Materials & Continua*, p. 11, 2018.

[32] Y. Lcun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[33] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 9409, pp. 94 090J–94 090J–10, 2015.

[34] G. Xu, H.-Z. Wu, and Y. Q. Shi, "Ensemble of CNNs for steganalysis: An empirical study," in *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*. New York, NY, USA: ACM, 2016, pp. 103–107.

[35] G. Xu, H. Z. Wu, and Y. Q. Shi, "Structural design of convolutional neural networks for steganalysis," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 708–712, 2016.

[36] J. Ni, J. Ye, and Y. I. Yang, "Deep learning hierarchical representations for image steganalysis," *IEEE Transactions on Information Forensics & Security*, vol. PP, no. 99, pp. 1–1, 2017.

[37] D. Hu, S. Qiang, S. Zhou, X. Liu, Y. Fan, and L. Wang, "Adaptive steganalysis based on selection region and combined convolutional neural networks," *Security & Communication Networks*, vol. 2017, no. 4, pp. 1–9, 2017.

[38] K. Zhong, G. Feng, L. Shen, and J. Luo, "Deep learning

for steganalysis based on filter diversity selection," *Science China(Information Sciences)*, vol. 61, no. 12, 2018.

[39] J. Zeng, S. Tan, B. Li, and J. Huang, "Large-scale JPEG image steganalysis using hybrid deep-learning framework," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1200–1214, 2018.

[40] G. Xu, "Deep convolutional neural network to detect J-UNIWARD," in *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2017, pp. 67–73.

[41] V. Mnih, N. Heess, A. Graves *et al.*, "Recurrent models of visual attention," in *Advances in neural information processing systems*, 2014, pp. 2204–2212.

[42] T. Xiao, Y. Xu, K. Yang, J. Zhang, Y. Peng, and Z. Zhang, "The application of two-level attention models in deep convolutional neural network for fine-grained image classification," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*. IEEE, 2015, pp. 842–850.

[43] X. Liu, T. Xia, J. Wang, and Y. Lin, "Fully convolutional attention localization networks: Efficient attention localization for fine-grained recognition," 2016.

[44] D. Zhao, Y. Chen, and L. Lv, "Deep reinforcement learning with visual attention for vehicle classification," *IEEE Transactions on Cognitive and Developmental Systems*, 2016.

[45] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," *arXiv preprint arXiv:1704.06904*, 2017.

[46] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.

[47] K. Kim, H. Lin, J. Y. Choi, and K. Choi, "A design framework for hierarchical ensemble of multiple feature extractors and multiple classifiers," *Pattern Recognition*, vol. 52, pp. 1–16, 2016.

[48] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot *et al.*, "Mastering the game of go with deep neural networks and tree search," *nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[49] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.*, "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, p. 354, 2017.

[50] W. Fei, M. Jiang, Q. Chen, S. Yang, L. Cheng, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," 2017.

[51] P. Bas, T. Filler, and T. Pevnỳ, "Break our steganographic system: The ins and outs of organizing BOSS," in *International Workshop on Information Hiding*. Springer, 2011, pp. 59–70.